

## BUNDESREPUBLIK DEUTSCHLAND

PRIORITY  
DOCUMENTSUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

DE00/2233



REC'D 25 SEP 2000

WIPO

PCT

4

Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung

**Aktenzeichen:** 199 30 570.6

**Anmeldetag:** 2. Juli 1999

**Anmelder/Inhaber:** Max-Planck-Gesellschaft zur Förderung der  
Wissenschaften e.V., München/DE

**Bezeichnung:** Pflanzen mit veränderter Genexpression

**IPC:** C 12 N 15/29

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 7. August 2000  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

Sellen

## Zusammenfassung



- Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung einer Pflanze mit veränderter Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen. Die vorliegende
- 10 Erfindung betrifft ferner ein Verfahren zur Herstellung von Pflanzen mit verändertem Flavonoidgehalt, umfassend das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

Neue deutsche Patentanmeldung  
Anmelder: Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V., Hofgartenstr. 8, 80539 München  
Titel: "Pflanzen mit veränderter Genexpression"  
Unser Zeichen: GI-001

2. Juli 1999

5

## Pflanzen mit veränderter Genexpression

Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung einer Pflanze mit veränderter  
10 Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen  
Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen  
Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt  
kodierenden Nukleinsäuresequenz in das Genom von Pflanzenzellen oder Pflanzengewebe und  
Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen. Die vorliegende  
Erfindung betrifft ferner ein Verfahren zur Herstellung von Pflanzen mit verändertem  
Flavonoidgehalt, umfassend das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß  
SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment  
oder Derivat in das Genom von Pflanzenzellen oder Pflanzengewebe und Regeneration der  
erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

20

Der pflanzliche Phenylpropanoidstoffwechsel, zusammenfassend beschrieben z.B. von  
Weisshaar & Jenkins, Current Opinion in Plant Biology 1, 251-257 (1998), oder von Shirley,  
Trends in Plant Sciences 1, 377-382 (1996), besteht aus einem Netz sich verzweigender  
biochemischer Reaktionsketten, die die Pflanze mit den verschiedensten phenolischen  
25 Komponenten versorgen. Der identische Beginn aller nachfolgenden Synthesen geht von  
Phenylalanin aus und führt über Cinnamat und 4-Coumarat zu Coumaroyl-CoA. Beteiligt sind  
die Enzyme Phenylalanin-Ammonium-Lyase (PAL), Cinnamat-4-Hydroxylase (C4H) und 4-  
Coumarat-CoA-Ligase (4CL). Im weiteren Verlauf verzweigen sich die Synthesen. Eine  
wichtige Verzweigung führt zur Biosynthese von Flavonoiden. Flavonoide sind Flavanderivate,  
30 d.h. ausschließlich bei Pflanzen vorkommende Sekundärmetaboliten mit dem zwei aromatische  
Ringe beinhaltenden Flavan-Grundgerüst. Zu den vielen, oft spezifisch in einer bestimmten  
Pflanzenart vorkommenden Flavonoiden zählen unter anderem UV-absorbierende Flavonole, zu  
den Gerbstoffen gehörende Tannine und z.B. als rote und blaue Blütenfarbstoffe gebildete  
Anthocyane.

35



In *Arabidopsis thaliana* sind verschiedene chromosomale Gen-Loci bekannt, die eine Rolle in der Flavonoidbiosynthese spielen. Mutationen in etlichen dieser Loci verhindern die Akkumulation brauner Farbstoffe in der Samenschale (Testa) und werden als transparent testa (tt in mutierter Form, TT als Wildtyp) bezeichnet. Durch die unter der Samenschale liegenden Kotyledonen  
5 erscheint der Samen in diesen Fällen gelblich bis hellbraun. Wildtypsamensamen sind dagegen dunkelbraun. Einige dieser Loci (tt3, tt4, tt5, ttg) sind außerdem noch an der Produktion von Anthocyanen in Blättern und Sproß beteiligt und ein Locus (ttg) hat eine zusätzliche Funktion bei der Entwicklung von Trichomen und Wurzelhaaren. Alle bisher aufgetretenen tt-Mutanten haben sich als rezessiv erwiesen. Eine zusammenfassende Beschreibung findet sich in Shirley et  
10 al., *The Plant Journal* 8, 659-671 (1995).

Die ersten *Arabidopsis*-Mutanten mit Defekt in der Flavonoidbiosynthese wurden 1971 von Bürger beschrieben (Bürger, *Arabidopsis Information Service* 8, 36-42 (1971)). In dieser Publikation wurde der Phänotyp von tt1, der eine abweichende Samenfarbe aufweist, erstmals  
15 erwähnt. Durch genetische und morphologische Untersuchungen von Koornneef (Koornneef, *Arabidopsis Information Service* 18, 45-51 (1981), Koornneef, *Arabidopsis Information Service* 27, 1-4 (1990)) sowie Shirley (Shirley et al., *The Plant Journal* 8, 659-671 (1995)) konnte der Genlocus von tt1 auf Chromosom 1 – 54,9 sowie die Beschränkung des Phänotyps auf den Samen festgestellt werden.

20 Bei vielen Nutz- und Zierpflanzen ist die Veränderung der Samenzusammensetzung und eine Veränderung der Flavonoidbiosynthese aus landwirtschaftlichem und produktionstechnischen Gesichtspunkten seit langem erwünscht. Eine begrenzte Einflußnahme auf Komponenten des Flavonoidstoffwechsels war, solange die Genstruktur der beteiligten Enzyme nicht bekannt war,  
25 nur mit den Mitteln der klassischen Züchtung möglich. Diese konventionelle Methode, beruhend auf der zufälligen Vermischung des väterlichen und mütterlichen Erbguts, ist relativ zeit- und kostenintensiv. Das gewünschte Ergebnis ist erst nach 10 bis 15 Jahren zu erwarten. Die gezielte Manipulation einzelner Komponenten der Flavonoidbiosynthese ist mit klassischer Züchtung ohne genetische Analysen kaum zu erreichen. Es besteht somit die Aufgabe, Pflanzen mit einer  
30 Veränderung der Zusammensetzung des pflanzlichen Samens und einer Verbesserung der Samenqualität von Pflanzen z.B. Nutz- und Zierpflanzen bereitzustellen.

Die Aufgabe wird durch den Gegenstand der Patentansprüche gelöst.

Die Erfindung wird durch die folgenden Figuren erläutert.

Figur 1 zeigt eine schematische Darstellung der Analyse der Nukleinsäuresequenz des TT1-Promoters durch Vergleich mit TRANSFAC MATRIX TABLE, Rel.3.3 (E. Wingender et al., 1998). Dabei sind Bindestellen für Transkriptionsfaktoren in Form von Großbuchstaben (SBF-1 like sites: siehe Lawton et al, Silencer region of a chalcone synthase promoter contains multiple binding sites for a factor, SBF-1, closely related to GT-1, Plant Molecular Biology 16, 235-249 (1991)), kursive Schrift (AGAMOUS like sites: Huang et al., Isolation and characterization of the binding sequences for the product of the Arabidopsis floral homeotic gene AGAMOUS, Nucleic Acids Research 21, 4769-4776 (1993)), unterstrichenen Buchstaben (P like sites: Grotewold et al., The myb-homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset, Cell 76, 543-553 (1994)), großen und kursiven Buchstaben (MYB Ph3 like sites: Solano et al., Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.PH3) from Petunia hybrida, EMBO Journal 14, 1773-1784 (1995)) sowie großen und unterstrichenen Buchstaben (Athab-1 und 2 like sites: Sessa et al., The athb-1 and-2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities, EMBO Journal 12, 3507-3517 (1993)) gekennzeichnet. Das Start-ATG ist fett und unterstrichen dargestellt. Die Numerierung beginnt mit der 5' gelegenen SpeI Schnittstelle im verwendeten Plasmidvektor pSK-TT1.

Figur 2 zeigt die Nukleinsäuresequenz der genomischen DNA-Sequenz von TT1, beginnend mit dem Start-ATG. Großbuchstaben stellen dabei Exons und kursiv geschriebene Buchstaben Introns dar. Die Numerierung setzt die Numerierung aus Figur 1 fort.

Figur 3 zeigt die für das TT1-Gen von Arabidopsis kodierende cDNA-Sequenz und die daraus abgeleitete Aminosäuresequenz von TT1.

Figur 4 zeigt schematisch einen Vergleich der TT1-Aminosäuresequenz mit Sequenzen aus der NCBI GenBank. Acc.No AL049660, AB025629 und AC006085.9 sind aus Nukleinsäuresequenzen abgeleitete hypothetische Aminosäuresequenzen aus Arabidopsis thaliana (At), AJ234704 ist eine aus einer Nukleinsäuresequenz abgeleitete hypothetische Aminosäuresequenz für Hordeum vulgare (Hv). In der Konsensussequenz bezeichnet ein ! Aminosäuren vom Typ I oder V, ein \$ Aminosäuren vom Typ L oder M, ein % Aminosäuren vom Typ F oder Y sowie ein # Aminosäuren vom Typ N, D, Q, E, B oder Z.

Figur 5 zeigt eine schematische Darstellung des Restriktionsmusters von pSK-TT1.

Figur 6 zeigt eine Darstellung der Samenfärbung der Mutante tt1 im Vergleich zum Wildtyp.

5

Figur 7 zeigt eine Darstellung der nukleäre Lokalisation von TT1.

Der hier verwendete Ausdruck „homologe Sequenz“ bezeichnet eine Nukleinsäure- oder Proteinsequenz mit signifikanter Ähnlichkeit zur Vergleichssequenz oder Teilen davon, wobei  
10 diese homologen Sequenzen eine Aktivität oder Teilaktivität vergleichbar der Aktivität der Nukleinsäure- oder Proteinsequenzen gemäß SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 oder SEQ ID NO:4 aufweisen. Als homologe Sequenzen gelten Nukleinsäuresequenzen, die mit den Sequenzen gemäß SEQ ID NO:1, SEQ ID NO:2 oder SEQ ID NO:3 oder Teilen dieser Sequenzen unter stringenten oder wenig stringenten Bedingungen hybridisieren (zu stringenten  
15 und wenig stringenten Bedingungen siehe Sambrook et al., Molecular Cloning, Cold Spring Harbour Laboratory (1989), ISBN 0-87969-309-6). Ein Beispiel für stringente Hybridisierungsbedingungen ist: Hybridisierung in 4 x SSC bei 65° C (alternativ in 50% Formamid und 4 X SSC bei 42° C), gefolgt von mehreren Waschschritten in 0,1 x SSC bei 65°C für insgesamt eine Stunde. Ein Beispiel für wenig stringente Hybridisierungsbedingungen ist  
20 Hybridisierung in 4 x SSC bei 37° C, gefolgt von mehreren Waschschritten in 1 x SSC bei Raumtemperatur. Als homologe Sequenzen sollen des weiteren Nukleinsäure- oder Proteinsequenzen oder Teile davon gelten, die unter Zuhilfenahme des Similaritätsalgorithmus BLAST (Basic Local Alignment Search Tool, Altschul et al., Journal of Molecular Biology 215, 403-410 (1990) eine signifikante Ähnlichkeit mit den Nukleinsäure- und Aminosäuresequenzen  
25 der vorliegenden Erfindung aufweisen. Als signifikant ähnlich werden, wie hier verwendet, Sequenzen bezeichnet, die z.B. unter Verwendung von Standardparametern im Blast-Service des NCBI ein Signifikanzniveau (Probability) von  $P < 10^{-5}$  aufweisen, wenn Sie mit den Sequenzen gemäß SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 oder SEQ ID NO:4 oder Teilen davon verglichen werden.

30

Der hier verwendete Ausdruck "Derivate" bezeichnet Nukleinsäuresequenzen, die eine oder mehrere Deletionen, Substitutionen, Additionen, Insertionen und/oder Inversionen aufweisen.

Der hier verwendete Ausdruck "funktional verbunden" bedeutet, daß eine regulatorische Sequenz wie ein Promotor die Expression eines Gens steuert oder daß eine Nukleinsäuresequenz von dem Promotor ausgehend exprimiert wird.

- 5 Der hier verwendete Ausdruck "Vektor" bezeichnet natürlich vorkommende oder künstlich erschaffene Konstrukte zur Aufnahme, Vermehrung, Expression oder Übertragung von Nukleinsäuren, z.B. Plasmide, Phagemide, Cosmide, künstliche Chromosomen, Bakteriophagen, Viren, Retroviren.
- 10 Der hier verwendete Ausdruck "Expressionssystem" bezeichnet jedwede Kombination von Vektoren, Restriktionsenzymen, Transformationsmethoden, Zellextrakten, lebenden Zellen z.B. prokaryotischen oder eukaryotischen Zellen oder Organismen mit dem Zweck der endogenen oder exogenen Expression von Genen.
- 15 Die vorliegende Erfindung betrifft die regulatorische samenspezifische Nukleinsäuresequenz (im folgenden auch Promotor bezeichnet), die natürlicherweise in *Arabidopsis thaliana* die Expression des TT1-Gens steuert. Die erfindungsgemäße Nukleinsäuresequenz ist in SEQ ID NO:1 aufgeführt. Ferner betrifft die Erfindung ein Fragment oder Derivat der Nukleinsäuresequenz gemäß SEQ ID NO:1 oder eine Nukleinsäuresequenz, die mit der
- 20 Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert und für die samenspezifische Expression verantwortlich ist. Bevorzugt ist eine Nukleinsäuresequenz, die unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert.
- 25 Die erfindungsgemäße Nukleinsäuresequenz kann natürlichen Ursprungs sein oder künstlich hergestellt worden sein. Die Nukleinsäuresequenz kann sowohl in Sense- als auch in Antisense-Orientierung vorliegen.
- 30 Die erfindungsgemäße Nukleinsäuresequenz eignet sich z.B. zur Identifizierung und Isolierung von zum TT1-Gen homologer Genen in anderen Organismen oder von homologen Genen in *Arabidopsis thaliana* mit Hilfe spezieller Hybridisierungs- oder Screening-Verfahren, z.B. als Sonde für das Screening in DNA-Bibliotheken mit Hilfe der Hybridisierung an einzelsträngige Nukleinsäuren ähnlicher Basenabfolge.

Die erfindungsgemäße Nukleinsäuresequenz eignet sich ferner zur spezifischen Steuerung der Expression von Genen in Organismen oder Zellen, bevorzugt zur spezifischen Steuerung der Expression von Genen in Samen, insbesondere in der Samenschale. Bevorzugt betrifft die vorliegende Erfindung ein Verfahren zur Herstellung einer Pflanze mit veränderter Genexpression, umfassend das stabile Integrieren einer samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat funktional verbundenen mit einer für ein Genprodukt codierenden Nukleinsäuresequenz in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen. Die mit der erfindungsgemäßen Nukleinsäuresequenz funktional verbundenen Nukleinsäuren können endogene, exogene genomische DNA-Abschnitte oder cDNAs oder deren Fragmente oder Derivate sein. Endogen bedeutet dabei, das die Nukleinsäuresequenz aus dem gleichen Organismus stammt, in den sie mit dem erfindungsgemäßen Verfahren integriert wird, z.B. eine Nukleinsäuresequenz aus *Arabidopsis thaliana* wird mit dem erfindungsgemäßen Verfahren in *Arabidopsis thaliana* integriert. Exogen bedeutet, das die Nukleinsäuresequenz aus einem anderen Organismus stammt, z.B. eine Nukleinsäuresequenz aus *Arabidopsis thaliana* wird mit dem erfindungsgemäßen Verfahren in z.B. Weizen integriert. Die Nukleinsäuresequenzen können gegenüber den natürlich vorkommenden Nukleinsäuresequenzen Deletionen, Substitutionen, Additionen, Insertionen und/oder Inversionen aufweisen.

Die erfindungsgemäße Nukleinsäuresequenz kann in Vektoren, Expressionssystemen oder Pflanzen, Pflanzengewebe oder Pflanzenzellen oder tierischen Zellen oder Mikroorganismen zur Veränderung der Expressionsmuster verschiedenster Genprodukte, verwendet werden. Die Expression der Genprodukte kann dabei gegenüber ihrer natürlichen Expression sowohl verstärkt als auch verringert sein.

Beispielsweise kann die erfindungsgemäße Nukleinsäuresequenz für die Expression des samenschalenspezifischen TT1-Gens verwendet werden. Ferner eignet sich die erfindungsgemäße Nukleinsäuresequenz auch für die Regulation der Expression anderer Gensequenzen für jedwede Anwendung, sowohl aus *Arabidopsis thaliana* als auch aus anderen Organismen. Dabei kann der Promotor in Kombination mit beliebigen Genen vorliegen, sowohl in einem Vektor als auch in transgenen Organismen.



Insbesondere kann die erfindungsgemäße Nukleinsäuresequenz zur Kontrolle der Expression weiterer natürlicher samenspezifischer oder künstlich in den Samen transferierter Gene eingesetzt werden, z.B. zur Expressionsregulation weiterer Gene des Phenylpropanoidstoffwechsels, z.B. Phenylalanin-Ammonium-Lyase, Cinnamat-4-Hydroxylase, 4-Coumarat-CoA-Ligase, Chalcon-Synthase, Chalcon-Isomerase, Chalcon-Reduktase, Flavanon 3-Hydroxylase, Flavonoid-3'-Hydroxylase, Flavonoid-3'5'-Hydroxylase, Dihydroflavonol-4-Reduktase, Leucoanthocyanidin-Dioxygenase, 3'-Glucosyltransferase, 5'-Glucosyltransferase, O-Methyl-Transferase. Die Kontrolle der Expression von Genen des Phenylpropanoidstoffwechsels mit Hilfe der erfindungsgemäßen Nukleinsäuresequenz eignet sich dabei insbesondere z.B. zur Verstärkung der UV-Absorptionsrate, Veränderung der Farbe, Verbesserung des Geschmacks oder der Lagerfähigkeit, Verstärkung des Schutzes vor Schädlingen oder Verbesserung der Verarbeitungsfähigkeit verschiedener Pflanzengewebe.

Ferner kann der die erfindungsgemäße Nukleinsäuresequenz zur Expressionsregulation von Genen, kodierend für weitere Proteine der Samenschale, verwendet werden. Beispiele für Proteine der Samenschale sind insektizid wirksame  $\alpha$ -Amylase-Inhibitoren und Proteinase-Inhibitoren und Faserproteine.

Außerdem ist die erfindungsgemäße Nukleinsäuresequenz zur Regulation der Speicherung von Reservestoffen im Samen, z.B. von Stärke, geeignet. Möglichkeiten der Einflußnahme auf Reservestoffe bestehen z.B. durch die Expression von Sense- oder Antisense-Transkripten des Kohlenhydratstoffwechsels der Pflanze wie z.B. ADP-Glucose-Synthetase, Stärkesynthase, ADP-Glucose-Pyrophosphorylase oder die Expression von Genen anderer Organismen wie z.B. Hefe-Invertase zur Mobilisierung der Stärke.

Die erfindungsgemäße Nukleinsäuresequenz eignet sich ferner zur Verlagerung von Stoffwechselprodukten des Samens, z. B. von Glucosinolaten zur Verbesserung der Nematodenresistenz, in die Samenschale. Ferner kann der Promotor zur Verhinderung oder Verzögerung der Reifung der Samenschale durch kontrollierte Expression von Ribonuklease-Genen benutzt werden. Eine Verzögerung der Reifung der Samenschale kann zur Bildung von größeren Samen führen. Des weiteren kann die Entfernung der Samenschale durch eine verzögerte Reifung erleichtert werden.

Geeignete Vektoren zur Aufnahme und Überführung der erfindungsgemäßen Nukleinsäuresequenz und der mit ihr funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz können die Vermehrung und/oder die Expression der aufgenommenen Nukleinsäuren in Einzellern wie z.B. *Escherichia coli* oder *Agrobacterium tumefaciens* oder in Pflanzenzellen, Pflanzengeweben oder Pflanzen oder tierischen Zellen oder Tieren gewährleisten. Entsprechende Vektoren können natürlich vorkommen oder aber künstlich hergestellt sein. Die Vektoren können Selektionsmarker, Terminatorsequenzen, Polylinker, Promotorelemente, Enhancer, Polyadenylierungsstellen und andere genetische Elemente umfassen. Zur Klonierung geeignete Vektoren sind z.B. pBluescript, Plasmide der pUC-Serie, Plasmide der pGem-Reihe oder auf dem Bakteriophagen  $\lambda$  basierende Vektoren. Ein zur Verwendung in *Agrobacterium* benutzter Plasmidvektor ist z.B. pBin19 (Bevan et al., *Nucleic Acids Research* 12, 8711-8721. (1984)). Zur Transformation und Expression in Pflanzen stellen auf dem Ti-Plasmid von *Agrobacterium*-Arten oder auf Pflanzenviren aufbauende Konstrukte verwendungsfähige Vektoren dar und sind dem Fachmann bekannt. Eine zusammenfassende Beschreibung bislang benutzter Vektoren findet sich in Guerineau und Mullineaux, *Plant Transformation and Expression Vectors*, in: *Plant Molecular Biology Labfax*, herausgegeben von Croy, Oxford, BIOS Scientific Publishers, 121-148 (1993).

Einige der in handelsüblichen Transformations- und Expressionssystemen angewendeten Transformationsmethoden zur Übertragung von Fremdgenen (Transformation) in das Genom von Pflanzen werden im folgenden vorgestellt. Die Wahl der Methode zur Einbringung der erfindungsgemäßen Nukleinsäuresequenz und der mit ihr funktional verbundenen für ein Genprodukt kodierenden Nukleinsäuresequenzen in pflanzliche Zellen ist jedoch nicht auf diese Liste beschränkt. Bislang eingesetzte Transformationsverfahren bei Pflanzen sind z.B. der Gentransfer mittels *Agrobacterium tumefaciens* (z.B. durch Baden von Samen oder Blattstückchen in einer Agrobakterienlösung), mittels pflanzlicher Viren, durch Elektroporation, durch Einschießen (microprojectile bombardment) oder Einspritzen (Mikroinjektion) sowie die Inkubation von trockenen Embryonen in DNA-haltigen Flüssigkeiten und die Transformation von Protoplasten unter Zuhilfenahme von Polyethylenglykol. Genauere Beschreibungen der angesprochenen Verfahren finden sich z.B. in Jens et al., *Techniques for Gene Transfer*, in: *Transgenic Plants*, Vol. 1, Engineering and Utilization, herausgegeben von Kung und Wu, Academic Press 128-143 (1993).

Die erfindungsgemäße Nukleinsäuresequenz kann zur Kontrolle der Genexpression in Mikroorganismen wie z.B. Escherichia coli oder Saccharomyes cerevisiae, in mono- und dikotylen Pflanzen sowie Algen und tierischen Zellen und Tieren verwendet werden. Besonders bevorzugte Pflanzen sind dabei Kulturpflanzen wie z.B., Soja, Reis, Baumwolle, Zuckerrübe, Canola, Sonnenblume, Flachs, Hanf, Kartoffel, Tabak, Tomate, Raps, Alfalfa, Salat, Erbse, Bohne, Karotte, Zwiebel, die verschiedenen Baum, Nuß- und Weinspezies, ferner Getreide wie z.B. Gerste, Weizen, Roggen, Hafer, Mais, ferner Obstsorten wie z.B. Mango, Apfel, Pfirsich, Stachelbeere, Johannisbeere, Banane, Melone, Kürbis und diverse Zitrusfrüchte wie z.B. Zitrone, Apfelsine, Pampelmuse, Mandarine. Die mit der erfindungsgemäßen Nukleinsäuresequenz transformierten Pflanzen können unmodifizierte Wildtyp-Pflanzen oder durch Züchtung erhaltene Pflanzen oder modifizierte Pflanzen z.B. transgene Pflanzen sein. Die erfindungsgemäße Nukleinsäuresequenz kann ferner nicht nur in Pflanzen oder Pflanzengeweben, sondern auch in einzelnen Pflanzenzellen, z.B. in einer Zellkultur, verwendet werden.

15

Die vorliegende Erfindung betrifft ferner eine transformierte Zelle, insbesondere eine transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, in der die erfindungsgemäße Nukleinsäuresequenz und die damit funktional verbundene für ein Genprodukt kodierenden Nukleinsäuresequenz stabil integriert sind. Ferner betrifft die vorliegende Erfindung eine mit der erfindungsgemäßen Nukleinsäuresequenz transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, die oder das zu einer samenproduzierenden Pflanze regenerierbar ist. Insbesondere betrifft die vorliegende Erfindung eine Pflanze, die nach dem erfindungsgemäßen Verfahren erhältlich ist. Ferner betrifft die vorliegende Erfindung Saatgut, das von Pflanzen erhalten wird, die nach dem erfindungsgemäßen Verfahren erhalten werden.

25

Die vorliegende Erfindung betrifft ebenfalls die genomische Sequenz oder die cDNA-Sequenz des TF1-Gens von Arabidopsis thaliana Ecotyp Columbia, das über die Bildung von Zwischenprodukten für die Bildung von Flavonoiden verantwortlich ist. Die erfindungsgemäße Nukleinsäuresequenz ist in SEQ ID NO:2 und 4 aufgeführt. Ferner betrifft die Erfindung ein Fragment oder Derivat der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert und für die Bildung von Flavonoiden verantwortlich ist. Bevorzugt ist eine Nukleinsäuresequenz,

30



die unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert. Die vorliegende Erfindung betrifft ferner die Aminosäuresequenz des TT1-Gens von *Arabidopsis thaliana* Ecotyp Columbia gemäß SEQ ID NO:3.

- 5 Die erfindungsgemäße Nukleinsäuresequenz eignet sich z.B. zur Identifizierung und Isolierung zum TT1-Gen homologer Gene in anderen Organismen oder von homologen Genen in *Arabidopsis thaliana* mit Hilfe spezieller Hybridisierungs- oder Screening-Verfahren, z.B. als Sonde für das Screening in DNA-Bibliotheken mit Hilfe der Hybridisierung an einzelsträngige Nukleinsäuren ähnlicher Basenabfolge.

10

Die vorliegende Erfindung betrifft ferner ein Verfahren zur Herstellung Organismen oder Zellen, insbesondere von Pflanzen, mit verändertem Flavonoidgehalt, umfassend das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat in das Genom von Zellen, insbesondere von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengewebe zu Pflanzen.

15

Die erfindungsgemäße Nukleinsäuresequenz oder deren Fragment oder Derivat kann natürlichen Ursprungs sein oder aber künstlich hergestellt worden sein. Die Nukleinsäuresequenz oder deren Fragment oder Derivat kann sowohl in Sense- als auch in Antisense-Orientierung verwendet werden. Ferner kann die erfindungsgemäße Nukleinsäuresequenz oder deren Fragment oder Derivat in den genomischen Locus des TT1-Gens oder eines zum TT1-Gen homologen Gens von *Arabidopsis thaliana* oder in einen genomischen Locus eines zum TT1-Gen homologen Gens einer anderen Pflanze durch homologe Rekombination integriert sein. Des weiteren kann die Nukleinsäuresequenz auch in Form von Ribonukleinsäuren z.B. als Ribozym verwendet werden. In diesem Fall sind die Thymin-Basen (T) durch Uracil-Basen (U) ersetzt. Die Bildung von Flavonoiden kann durch die erfindungsgemäße Nukleinsäuresequenz sowohl verstärkt als auch verringert sein.

20

- 30 Die erfindungsgemäße Nukleinsäuresequenz kann z.B. zur Expression in Mikroorganismen z.B. *Escherichia coli* oder *Saccharomyces cerevisiae* und in mono- und dikotylen Pflanzen sowie Algen und tierischen Zellen und Tieren verwendet werden. Bevorzugte Pflanzen sind dabei Kulturpflanzen z.B., Soja, Reis, Baumwolle, Zuckerrübe, Canola, Sonnenblume, Flachs, Hanf, Kartoffel, Tabak, Tomate, Raps, Alfalfa, Salat, Erbse, Bohne, Karotte, Zwiebel, die

verschiedenen Baum, Nuß- und Weinspezies sowie Getreide z.B. Gerste, Weizen, Roggen, Hafer, Mais sowie Obstsorten z.B. Mango, Apfel, Pfirsich, Stachelbeere, Johannisbeere, Banane, Melone, Kürbis und diverse Zitrusfrüchte z.B. Zitrone, Apfelsine, Pampelmuse, Mandarine. Die mit der erfindungsgemäßen Nukleinsäuresequenz transformierten Pflanzen können  
5 unmodifizierte Wildtyp-Pflanzen oder durch Züchtung erhaltene Pflanzen oder modifizierte Pflanzen z.B. transgene Pflanzen sein. Die erfindungsgemäße Nukleinsäuresequenz kann ferner nicht nur in Pflanzen oder Pflanzengeweben sondern auch in einzelnen Pflanzenzellen z.B. in einer Zellkultur verwendet werden.

- 10 Eine Expression der erfindungsgemäßen Nukleinsäuresequenz kann durch Kombination der Sequenz mit einem geeigneten Promotor erreicht werden. Der in dieser Kombination vorliegende Promotor kann dabei sowohl der TT1-Promotor gemäß SEQ ID NO:1 als auch ein anderer endogener Promotor der transformierten Zelle oder ein auf dem Vektor befindlicher exogener Promotor sein. Als Promotor ist dabei grundsätzlich jede regulatorische Sequenz geeignet, die  
15 die Expression von Fremdgenen in Zellen, insbesondere in Pflanzen steuern kann, z.B. der CaMV 35S-Promotor aus dem Blumenkohl-Mosaik-Virus (Franck et al., Cell 21, 285-294 (1980)). Eine Expression der erfindungsgemäßen Nukleinsäuresequenz kann auch durch einen chemisch induzierbaren Promotor erreicht werden. Beispiele für chemisch induzierbare Promotoren sind der PRPI-Promotor (Ward et al., Plant Molecular Biology 22, 361-366 (1993)),  
20 ein durch Salizylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzenesulfonamid induzierbarer Promotor (EP-A 388186), ein durch Tetrazyklin induzierbarer Promotor (Gatz et al., Plant Journal 2, 397-404 (1992)), ein durch Abscisinsäure induzierbarer Promotor (EP-A 335528) sowie ein durch Ethanol oder Cyclohexanon induzierbarer Promotor (WO 93/21334). Je nach gewünschtem Expressionsort können auch Promotoren verwendet werden, die in  
25 bestimmten Pflanzengeweben oder Pflanzenteilen aktiv sind. Beispiele für entsprechende Promotoren sind der Phaseolin-Promotor (US 5504200), der Isoflavon-Reduktase Promotor (US 5750399), ein samenspezifischer Promotor aus Tabak (US 5824863) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al., EMBO Journal 8, 2445-2452 (1989)).
- 30 Ferner ist die erfindungsgemäße Nukleinsäuresequenz kombinierbar mit Sequenzen, die ein Targeting in bestimmte Kompartimente der Pflanze sicherstellen, z.B. für Transitpeptide oder Teile davon kodierende Sequenzen. Des weiteren sind Sequenzen, kodierend für enzymatisch aktive oder antigen wirksame Proteine z.B. His-tag mit der obengenannten erfindungsgemäßen Nukleinsäuresequenz kombinierbar.

Ferner kann die erfindungsgemäße Nukleinsäuresequenz in Vektoren, Expressionssystemen oder Pflanzen, Pflanzengeweben oder Pflanzenzellen oder tierischen Zellen oder Mikroorganismen zur Veränderung des Expressionsmusters verwendet werden.

- 5 Die erfindungsgemäße Nukleinsäuresequenz eignet sich ferner in Kombination mit verschiedenen Promotoren zur Manipulation der phänotypischen und genotypischen Eigenschaften verschiedener Pflanzen oder Pflanzengewebe, z.B. zur Veränderung der Samenfarbe, z.B. zur ästhetischen Verbesserung verschiedener Pflanzenarten. Zierpflanzen mit z.B. ausgeschaltetem oder mutiertem TT1-Gen können optisch attraktive Varietäten darstellen.

10

Die erfindungsgemäße Nukleinsäuresequenz kann ebenso zur Verstärkung der UV-Schutzfunktion der Samenschale durch Veränderung des Flavonoidgehaltes verwendet werden. Flavonoide zeigen Absorptionsmaxima im ultravioletten Bereich. Absorbierend wirken dabei die delokalisierten  $\pi$ -Elektronen der phenolischen Gruppen. Eine Erhöhung der

15 Flavonoidkonzentration im Samen kann eine drastische Verringerung der UV-induzierten Gewebeschäden bewirken. Auf diese Weise läßt sich z.B. die Keimungsrate des Saatgutes, vor allem in Regionen mit intensiver Sonnenbestrahlung, verbessern.

- Ferner kann die erfindungsgemäße Nukleinsäuresequenz zur Verstärkung der Schutzfunktion der
- 20 Samenschale gegen Schädlingsbefall durch Veränderung des Flavonoidgehaltes verwendet werden. Proanthocyanidine und andere phenolische Komponenten wirken fungizid, u.a. aufgrund enzyminhibitorischer Eigenschaften (Jambunathan et al., Journal of Agricultural and Food Chemistry 34, 425-429 (1986)). Eine Konzentrierung dieser Stoffe in der Samenschale kann die Pathogenresistenz erhöhen.

25

- Ferner kann die erfindungsgemäße Nukleinsäuresequenz zur Verbesserung der Geschmacksqualität des Samens und anderer Pflanzenteile infolge einer Veränderung des Flavonoidgehaltes, insbesondere des Gehaltes an kondensierten Tanninen, verwendet werden. Kondensierte Tannine haben einen großen Einfluß auf den Geschmack vieler Obst- und
- 30 Gemüsesorten z.B. von Apfel, Kiwi oder Banane. Auf pflanzlichen Extrakten basierende Getränke z.B. Kaffee, Tee, Wein und Fruchtsäften werden ebenfalls in ihrer Geschmacksqualität von Tanninen geprägt.

Eine weitere Verwendung der erfindungsgemäßen Nukleinsäuresequenz stellt die Verbesserung der Verarbeitungsfähigkeit des Samens in industriellen Produktionsprozessen dar. Beispielsweise verursachen kondensierte Tannine in der Testa des Samens der Gerste unerwünschte Präzipitate während des Bierbrauprozesses (Shirley, Seed Science Research 8, 415-422 (1998)). Durch  
 5 reduzierte Tanningehalte kann die Präzipitatbildung verhindert werden.

Die vorliegende Erfindung betrifft ferner eine transformierte Zelle, insbesondere eine transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, in der die erfindungsgemäße Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen  
 10 Nukleinsäuresequenz, oder deren Fragment oder Derivat stabil integriert ist. Ferner betrifft die vorliegende Erfindung eine mit der erfindungsgemäßen Nukleinsäuresequenz transformierte Pflanzenzelle oder ein transformiertes Pflanzengewebe, die oder das zu einer samenproduzierenden Pflanze regenerierbar ist. Insbesondere betrifft die vorliegende Erfindung eine Pflanze, die nach dem erfindungsgemäßen Verfahren erhältlich ist. Ferner betrifft die  
 15 vorliegende Erfindung Saatgut, das von Pflanzen erhalten wird, die nach dem erfindungsgemäßen Verfahren erhalten werden.

## BEISPIELE

20

Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt:

### Beispiel 1: Allgemeine Klonierungsverfahren

25

Die im Rahmen der vorliegenden Erfindung durchgeführten Klonierungsschritte wie z.B. Restriktionsspaltungen, Agarose-Gelelektrophorese, Reinigung von Nukleinsäurefragmenten, Transfer von Nukleinsäuren auf Filtermaterialien, Transformation und Anzucht von Bakterienzellen u.s.w. wurden wie bei Sambrook et al., Molecular Cloning, Cold Spring Harbour Laboratory (1989), ISBN 0-87969-309-6, durchgeführt.

30

### Beispiel 2: Herstellung einer Knockout-Population von Arabidopsis thaliana

Die vorliegende Erfindung wurde durch das Screening einer mit dem Transposon En-1/Spm (Pereira et al, EMBO Journal 5, 835-841 (1986)) mutagenisierten Knockout-Population von Pflanzen der Spezies Arabidopsis thaliana Ecotyp Columbia erhalten. Die Integration der

transposablen Elemente in Gene der Mutterpflanze führt häufig zum Ausfall der entsprechenden Genfunktion und in vielen Fällen zu einem vom Wildtyp unterscheidbaren Erscheinungsbild der betroffenen Pflanze. Zum Aufbau der Knockout-Population wurde das autonome En-1 Element aus *Zea mays* mittels *Agrobacterium tumefaciens* in *Arabidopsis* übertragen. Das entsprechende Transposon tagging System ist in Cardon et al., *Plant Molecular Biology* 23, 157-178 (1993) beschrieben. Das verwendete Ti-Plasmid, pGV3850HPT::pkEn2, beinhaltet das komplette En-1 Element als Integrat. Zur Selektion von Hygromycin-resistenten Transformanten trägt dieser Vektor das HPT-Gen unter der Kontrolle des viralen CaMV 35S Promotors. Samen einer Transformante mit einer einzelnen T-DNA-Insertion wurden auf Hygromycin-haltigem Medium ausgesät. Samen der so entstandenen, gegen Hygromycin resistenten Pflanzen (T<sub>2</sub>-Generation) wurden anschließend auf Kanamycin-haltigem Medium ausgesät. Bei den auf diese Weise selektierten Pflanzen (T<sub>3</sub>-Generation) war das En1-Element aus der T-DNA heraus transponiert. Mittels PCR wurde diejenigen Pflanzen der T<sub>4</sub>-Generation identifiziert, die ein oder mehrere transponierte En-1 Elemente, jedoch keine En-1 Elemente mehr in der T-DNA trugen. Diese Pflanzen beinhalteten jedoch noch die T-DNAs ohne integrierte En-1 Elemente. Zur Erzeugung von En-1 positiven, T-DNA negativen Pflanzen wurde die T<sub>4</sub>-Generation mit dem Wildtyp *Arabidopsis thaliana* Ecotyp Columbia gekreuzt und En-1 positive, T-DNA negative Pflanzen (S<sub>0</sub>-Generation) durch PCR identifiziert. Samen jeder dieser Pflanzen wurden über 6 - 12 Generationen (bis zur S<sub>6</sub>- bzw. S<sub>12</sub>-Generation) hinweg vermehrt, bis insgesamt 3 000 Linien mit insgesamt 15 000 unabhängigen En-1 Insertionen zur Verfügung standen (Wisman et al., *Plant Molecular Biology* 37, 989-999 (1998), Baumann et al., *Theoretical and Applied Genetics* 97, 729-734 (1998)).

### Beispiel 3: Screening nach TT-Mutanten

Zur Identifizierung von phänotypisch auffälligen Mutanten der entstandenen En-1 Population wurden 2000 Familien der S<sub>6</sub>- Generation (mit jeweils 20 Individuen) per Auge nach Auffälligkeiten durchmustert. Dabei konnte eine Linie (5K69) identifiziert werden, die sich durch eine abweichende Farbe der Samenschale (gelb statt dunkelbraun) auszeichnete, während die Produktion von Anthocyanen in Sproß und Blättern augenscheinlich nicht beeinträchtigt war.

Die Beschränkung des Phänotyps auf den Samen wurde von Koornneef, supra, und Shirley, supra, auch für die klassische *tt1*-Mutante beschrieben. Um zu überprüfen, ob es sich bei der in der En-Population gefundenen Linie um ein Allel von *tt1* handelte, wurden beide Linien



miteinander gekreuzt. Die Nachkommen dieser Kreuzung produzierten wiederum gelbe Samen. Das deutete darauf hin, daß tatsächlich beide Eltern ein defektes Allel des TT1 Gens tragen und vererbt haben.

- 5 Zur Klonierung eines Gens wurden DNA-Abschnitte bestimmt, die eine bekannte DNA-Sequenz, z.B. ein Transposon, flankieren. Dazu mußte zunächst gezeigt werden, daß sich in der Linie 5K60 das En-Transposon immer noch im TT1 Gen befand. Für diesen Zweck wurde eine Population von 51 Schwesterpflanzen der tt1-En Linie mittels Southern Blotting analysiert. 19 dieser Pflanzen produzierten gelbe und 33 dunkelbraune Samen. Es konnten verschiedene
- 10 Banden identifiziert werden, die mit einer En-Sonde hybridisierten. Eine dieser Banden fand sich in allen Pflanzen, die gelbe Samen hervorbrachten, sowie in 16 der braunsamigen. Sie fehlte in den übrigen 17 Pflanzen. Dieses Ergebnis bestätigt die Annahme, daß alle gelbe Samen
- 15 produzierende Pflanzen homozygot für eine Insertion des En-Transposon am tt1 Locus sind, während die braunsamigen Pflanzen entweder heterozygot für diese Insertion oder homozygot für den Wildtyp sind.

Die flankierende DNA dieser Insertion wurde durch schnelle Amplifikation genomischer Enden (RAGE), vgl. Cormack und Somssich, Gene 194, 273-276 (1997), gewonnen und in den pCR-TOPO Vektor (Firma Invitrogen) kloniert. Das resultierende Plasmid ist pCR-RAGE. Das Insert

20 von pCR-RAGE wurde als Sonde verwendet, um die IGF-BAC Bücherei (Mozo et al., Plant Journal 16, 377-384 (1998)) zu durchsuchen. Dabei wurden 5 positive Klone identifiziert (11O6, 3N5, 2B22, 10P4, 4M12). Alle diese Klone sind im Arabidopsis thaliana Genom auf Chr.I bei etwa 55cM lokalisiert, was mit der Kartenposition der tt1-Mutation übereinstimmt.

#### 25 Beispiel 4: Sequenzierung der genomischen TT1-Region

Ein 12 kb großes SpeI-Fragment des BACs 3N5, das mit der Sonde hybridisiert, wurde in pSK Bluescript (Firma Stratagene) subkloniert. Das resultierende Plasmid ist pSK-TT1. pSK-TT1 wurde unter Verwendung des Genome Priming Systems GPS1 der Firma New England Biolabs und eines Sequenzierautomaten der Firma ABI, Modell 377, durchsequenziert. Die Sequenzen

30 wurden zusammengefügt und mit verschiedenen Computerprogrammen untersucht. Vergleiche von TT1 mit Sequenzen aus der NCBI GenBank ergaben Ähnlichkeiten mit Zinkfingerproteinen. Zinkfingerproteine sind in der Lage, an DNA-Sequenzen zu binden und regulatorische Funktionen bezüglich der Expression bestimmter Gene auszuüben. Die von der TT1-cDNA abgeleitete Proteinsequenz von 303 AS Länge zeigt Ähnlichkeiten von etwas über 30% zu

Zinkfingerproteinen wie StPCP1 (X82328, Kühn und Frommer 1995, MGG 247, 759-763) und ZmID1 (AF0058757, Colasanti et al, 1998, Cell 93, 593-603). Eine weitaus höhere Ähnlichkeit von über 70% besteht zu Datenbankeinträgen, die bislang allerdings nur hypothetische Proteine repräsentieren. Der Computervergleich zeigt neben der TT1-Aminosäuresequenz hypothetische Aminosäuresequenzen aus *Arabidopsis thaliana*, die aus den Sequenzen mit den Acc.No AL049660, AB025629 oder AC006085.9 und für *Hordeum vulgare* aus AJ2347041 abgeleitet wurden. Die Ähnlichkeiten erstrecken sich in diesem Fall über einen deutlich grösseren Bereich der Sequenz, gehen also über die Zinkfingerregion hinaus.

10

#### Beispiel 5: Ermittlung der TT1-cDNA

Die Existenz eines exprimierten Gens sowie die Position des putativen Introns wurden mittels RT-PCR überprüft. Durch 3' und 5'-Race wurde die Länge der vollständigen cDNA bestimmt.

15

#### Beispiel 6: Komplementation der tt1-Mutation

Um zu zeigen, daß es sich bei dem klonierten Gen tatsächlich um TT1 handelt, wurde die tt1 Mutation komplementiert. Dazu wurde das 12 kb Insert aus pSK-TT1 in die SpeI-Schnittstelle des Vektors pGPTV-Kan-TATA::GUS umklontiert. Dieser Vektor entstand aus pGPTV-Kan (Becker et al., Plant Molecular Biology 20, 1195-1197 (1992)) durch Austausch des CaMV 35S-Promotors gegen einen Polylinker. Nach Transformation in den Agrobakterienstamm GV3101 (mit dem Virulenzplasmid pMP90, Koncz und Schell, Molecular and General Genetics 204, 383-396 (1986)) wurden tt1-Pflanzen mittels Vakuuminfektion transformiert (Bechthold et al., Molecular Biology and Genetics 316, 1194-1199 (1993)). Transformanten wurden auf Kanamycinhaltigem MS-Medium selektiert und auf ihre Samenfarbe hin analysiert. Samen komplementierter Pflanzen entsprachen bezüglich ihrer Färbung dem Wildtyp.

25

#### Beispiel 7: Ermittlung der zellulären Lokalisation des TT1-Proteins

Um die zelluläre Lokalisation des TT1-Proteins zu ermitteln, wurde es c-terminal an das „Grün fluoreszierende Protein“ (GFP) fusioniert. Unter Verwendung von pAVA393 (von Arnim et al., Gene 221, 35-43 (1998)) und der kompletten cDNA entstand so pTT1-GFP. Das Plasmid wurde in *Arabidopsis*-Protoplasten transfiziert (Hartmann et al., Plant Molecular Biology 36, 741-54 (1998)) und nach zwanzigstündiger Inkubation die GFP-Fluoreszenz bestimmt. Anhand dieser Untersuchung konnte die Lokalisation des TT1-Proteins im Zellkern nachgewiesen werden.

30

### Beispiel 8: Expressionsanalysen

Zur Untersuchung des TT1-Promotors wurden 1 bzw 3 kb grosse Fragmente zunächst in den Vektor pBT10 (Feldbrügge et al., Plant Journal 11, 1079-1093 (1997)) vor GUS gesetzt und die ganze Promotor-GUS Kassette danach in den binären Vektor pGPTV überführt. Nach  
5 Transformation in *Agrobacterium* wurden damit Wildtyp-Pflanzen von *Arabidopsis thaliana* Columbia infiltriert.

## SEQUENZPROTOKOLL

<110> Max-Planck-Gesellschaft z. Förd. d. Wissenschaften

<120> Pflanzen mit veränderter Genexpression

<130> GI-001

<140> xx

<141> 1999-07-02

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 3389

<212> DNA

<213> Arabidopsis thaliana

<400> 1

```

actagttgac cacatgaact aaacttcttg, gacaatcatc aatggacaca tgtagcttt 60
gatttgcctgt gaatttgctt tatctctcag tataattatc acttctctgt ttatgcttac 120
25 aatatatttt atggttttaga gttttgtttt acgatttttg atttaatgga taaagattag 180
ggattgaggg tttgagttaa gggttaaggaa attaggcttt agtgtagagt ctcaaggggt 240
taaggtttac acaccacaaa ccatttgctt gtgtcaacaa cattgtatca tttttcaaa 300
aaaattttgt tgaaggacct tgtattgata tatataaagc gaactgtttg gataagttta 360
tgtggacaat atatatgtga tacataatta gaaacatagt ttaatatctg atattgttg 420
30 ggaatatata atactactta ggtttaaata tatagtattt catatgatgc gaactgtttg 480
gataagttta cgtggacaat atatatgtga tacataatta ggaacatagt ttaatatgtg 540
atatattgtt ggaatatata attctactta cgcttaaata tttttatttg aattaaagca 600
tttcatataa tgtgaactgt ttgaatatgt ttacatggac aatatatatt ggatacataa 660
ttaggaacat agtttaatat ctgatatttg ttggaaatat ataatatag ttaagcttaa 720
35 atatttttat ttgatataat atttgactta aacattttta ttgattaaa cttaaatttta 780
acagatctta ccattaattt ttaacttggt atctctatct aatgtcacgt atattgtttt 840
ttagtaattg gcaacaaaat taatttatct cctgtttttt ttccttctca cctttataag 900
ggtaaaatgg tcataaaatc agtaaaaaag gtgaaaaagt gccactccc tcaaaagtgt 960
cataaacgtc caaactttct ccataaatgc cttatttttg aacattccat atagattata 1020
40 acttattata gggtataact tattatagtt acgttaatta tatgaatttc tattagttat 1080
cacacaatca aatatattta tcacaaaaat ttattaaaca ttttatatgt gtagtataa 1140
tgcaataaca tattatatgt ggtggcataa tgcaacaaca tattatttgt ctacgaatct 1200
cctttatttt tcgtttatgt aacaacagta aaacggattg tttagcttga tattctatat 1260
tataataatc taaagttatt tttgtaaatt atttttttcc caaattggat aaccaatcat 1320
45 agagttggta tttttttttt ttgttaaaat atatatatac tgaatatcga gttattgtgc 1380
atgacaattt atatggcgga cgagtttaaa tcgacattaa taacaattaa aatattatta 1440
atctaatact taaatactgg ttaaatcacc aatttattat tcttaatacc acatattaaa 1500
catatctaatt ctttactgat tcaataaaga ttgtgtgaaa caaaagtgtg cttgcaaaga 1560
attaatattg tacatagatt ttgttctggt agctagtact aaaatccatt aataaaacta 1620
50 atacggtatc tttattgatc atgtaacatg aattattcat gtatatacaa ttgaccctat 1680
taattttgca taaatttcaa cttggcaaat tcattgattt tgtaaacggt taattctgct 1740
aatttcacaa ttctcttgta cgctaaaaat ttatgcgtat tatcgtattg atatgcaaat 1800
atcgaagaat ttatagtttt atatagtaga aatgaaggta ttgcaaaaac gagttctaac 1860
gtgaaataac actaattaat taattagagt ttgaacctac agagattcga cttgatccac 1920
55 ttgaaaaatt catttactct actaatttgg ttactccatg gaccatgatt atgctattct 1980
gtaggactct aacaactgac ttgacacaat ctctttcgtg aacaataatg gggtattatt 2040
ttttgttttg ttttttcgga caaattagcc acgttgcttt agaccatttt gtagttctta 2100
tcttgaatca aagttctcagc taaaaaaaac aaaaaaacgc ttaaattcac tagctagact 2160
acgactacgt tggttaaatg ttttttttta aatacaatac attgaagtta aatatttgaa 2220
60 taaagaaaaat ctaatcagca tgtatacagt atattagaag taatacttga tcagaaaaat 2280
aacatacaat aataaaaata taaaaaaatt atgttagttt ttgggaatat tataattcta 2340
ctttcaatca aaataactaa aagaaataaa atcttcacac atagtggtaa taattggcta 2400

```

gtatgaatat tgaattgtgg agaccgggca taatatttga ctaggcagaa attattgata 2460  
 tgtactaagt taataacctt gcaaagaaat tcttttagtg aaacgtgtac atttgtaaaa 2520  
 acagatttaa cactaaatct tgacttgtat atactattaa ttattccttt tctcttattg 2580  
 5 gtatgtcaaa tctagtgttt acaaaaccag aggtgttgac cgtagagag agaattaaac 2640  
 aactttacata catacaaaac tagcattcat ttcattacc aaatccgaaa ttctattgat 2700  
 taataatatg aattcaacat tgtagaaaag ctaaaagctt acgtaaatgc caaagatagt 2820  
 taaaattaat acaattgtat tgtagaaaag tcttgaagag ttgggtccac aaaatttaag 2880  
 10 caaaaccctg caatgacaaa gttgcaaaaa tcttgaagag ttgggtccac aaaatttaag 2940  
 gttcttgttt ttccactcta tttataggca aagagatgag acagagaaga ttaaattact 3000  
 tcttaacaaa gggtgttttc actcaaccac atgcattctc aagtgtctgc tctcacatt 3060  
 cccaagatt cccatttact cacttctcta tttggtacgt aagtcacaca atatgattct 3120  
 aaattatttt acacattatt cgttttgttc acacttgctt tgcactttcg taaacctata 3180  
 tagttcatcc aatattattc ggtaaattcg atatttatca atctttattc tcgtaggtta 3240  
 15 aaggagacga ttgatacgtg ggatctactt acgtatctgc atgattatta gttataaaaag 3300  
 ttattgcaaa cattaaatta ctttcataga gagcaatcat tatattaagg taatttaatt 3360  
 ttattatata tagtcaagat ttaaaggaat aaagaaaaga ttctcaaac atttcatctc 3389  
 tctccaacaa ctattcacca cattcaatg

20 <210> 2  
 <211> 912  
 <212> DNA  
 <213> Arabidopsis thaliana

25 <400> 2  
 atggagtcac caccactata cgagatatcc tcaagctctt cttctgaaaa acctagacac 60  
 catttccaat cccttgatct cttccctaac ctcaaccaaa actcttgtat caacaatacc 120  
 ctaattgagc ctttaccgct tattgatgcg ataaacttga actcaaacct agacctaaac 180  
 30 cctaattccct tgtatgcgga agaaggagag caagaggagg aagaagaaga agaagaagac 240  
 cgtgaagtgg acgtggactt acacatcggc cttcctggtt ttggtaaacc aagcaatgat 300  
 gctaaacagc tgaagaagag aaatgggaag gagatcgcca catatgacgc cggaaaaggc 360  
 atcgagaatg aactttccgg aaaggcatac tggatcccg cgccggagca aattctcata 420  
 gggttcactc atttttcttg ccattgtatgc ttcaagacat tcaatcgcta caacaatctt 480  
 35 cagatgcaca tgtggggaca tggttcacaa tacaggaaaag gaccggagtc actgaaaggc 540  
 acacagccac gagccatggt agggatccct tgttactgct gcgttgaagg gtgcaggaaac 600  
 cacattgacc atcctcgctc caagccactg aaagacttta ggacgctcca aacgcactac 660  
 aaacgcaaac acggacacaa acccttctcg tgcgccttt gcggttaagct tttggctgtc 720  
 aagggcgatt ggcaaacaca tgagaagaat tgtggaaaac gttgggtttg cgtttgcggt 780  
 tctgatttta aacacaaacg ttctcttaag gaccatgtta aggcgtttgg gtctggctcat 840  
 40 gggccttata caactgggtt gtttgaagag caggcttcta attcatctgt ctccgagact 900  
 ttgttttttt aa 912

<210> 3  
 45 <211> 303  
 <212> PRT  
 <213> Arabidopsis thaliana

M E  
 <400> 3 S P P L Y E I S S S S S S E  
 50 Met Glu Ser Pro Pro Leu Tyr Glu Ile Ser Ser Ser Ser Ser Ser Glu  
 1 5 10 15  
 Lys Pro Arg His His Phe Gln Ser Leu Asp Leu Phe Pro Asn Leu Asn  
 20 25 30  
 55 Gln Asn Ser Cys Ile Asn Asn Thr Leu Ile Glu Pro Leu Pro Leu Ile  
 35 40 45  
 60 Asp Arg Ile Asn Leu Asn Ser Asn Leu Asp Leu Asn Pro Asn Pro Leu  
 50 55 60  
 Tyr Ala Glu Glu Gly Glu Gln Glu Glu Glu Glu Glu Glu Glu Asp  
 65 70 75 80

Arg Glu Val Asp Val Asp Leu His Ile Gly Leu Pro Gly Phe Gly Lys  
85 90 95

5 Pro Ser Asn Asp Ala Lys Gln Leu Lys Lys Arg Asn Gly Lys Glu Ile  
100 105 110

Ala Thr Tyr Asp Ala Gly Lys Gly Ile Glu Asn Glu Leu Ser Gly Lys  
115 120 125

10 Ala Tyr Trp Ile Pro Ala Pro Glu Gln Ile Leu Ile Gly Phe Thr His  
130 135 140

Phe Ser Cys His Val Cys Phe Lys Thr Phe Asn Arg Tyr Asn Asn Leu  
145 150 155 160

15 Gln Met His Met Trp Gly His Gly Ser Gln Tyr Arg Lys Gly Pro Glu  
165 170 175

20 Ser Leu Lys Gly Thr Gln Pro Arg Ala Met Leu Gly Ile Pro Cys Tyr  
180 185 190

Cys Cys Val Glu Gly Cys Arg Asn His Ile Asp His Pro Arg Ser Lys  
195 200 205

25 Pro Leu Lys Asp Phe Arg Thr Leu Gln Thr His Tyr Lys Arg Lys His  
210 215 220

Gly His Lys Pro Phe Ser Cys Arg Leu Cys Gly Lys Leu Leu Ala Val  
225 230 235 240

Lys Gly Asp Trp Arg Thr His Glu Lys Asn Cys Gly Lys Arg Trp Val  
245 250 255

35 Cys Val Cys Gly Ser Asp Phe Lys His Lys Arg Ser Leu Lys Asp His  
260 265 270

Val Lys Ala Phe Gly Ser Gly His Gly Pro Tyr Pro Thr Gly Leu Phe  
275 280 285

40 E E R A S N S S V S E T L P A  
Glu Glu Gln Ala Ser Asn Ser Ser Val Ser Glu Thr Leu Phe Phe  
290 295 300

<210> 4  
<211> 1816  
<212> DNA  
<213> Arabidopsis thaliana

50

<400> 4  
atggagtcac caccactata cgagatatcc tcaagctctt cttctgaaaa acctagacac 60  
catttccaat cccttgatct cttccctaac ctcaaccaa actcttgat caacaatacc 120  
ctaattgagc ctttaccgct tattgatcgc ataaacttga actcaaacct agacctaaac 180  
55 cctaattccct tgtatgcgga agaaggagag caagaggagg aagaagaaga agaagaagac 240  
cgtgaagtgg acgtggactt acacatcggc cttcctgggt ttggtaaacc aagcaatgat 300  
gctaaacagc tgaagaagag aaatgggaag gagatcgcca catatgacgc cggaaaaggc 360  
atcgagaatg aactttccgg aaaggcatac tggatcccg cgccggagca aattctcata 420  
gggttcactc atttttcttg ccattgatgc ttcaagacat tcaatcgcta caacaatctt 480  
60 caggtagcag tcaatatatc tcatgcgcgt tgcttttcca tgcacaaaca tatataataa 540  
attcatctta tagagttata tctccggatc taatgttatg agttttattca tatctatata 600  
tatacatata tatatatata tatatatata attctgaatt tatttgataa 660  
aagctaaaca accaggattt aatagatgat ttacctttgg atcttattat acaatttaca 720

aatttaatca agtcaactaa tcgtgattta attacttttt tttgtaagaa gagttggttaa 780  
tatatatattt tatggtaatg ttttcatgaa aataattcat cacaactctt tacattttatt 840  
taatgcctta actaaagctg aattcgaaaa agttgaaata aattatctac taagatttga 900  
5 ttgactatag tttttaatag ttttcttttc tcatatatat attatcatag tagtcaaaac 960  
atgtgattca aacttaaata cacagatttc ttgaatgaaa cattactatg ctcggtcaat 1020  
aatatgattt taaggaaacca tgttatttca ttttattact taaggaaacc tttttgtttt 1080  
ttgttgactc taaatattat gaatatagat gcacatgtgg ggacatggtt cacaatacag 1140  
gaaaggaccg gagtcaactga aaggcacaca gccacgagcc atgttaggga tcccttggtta 1200  
10 ctgctgcgtt gaaggggtgca ggaaccacat tgaccatcct cgttccaagc cactgaaaga 1260  
cttttaggacg ctccaaacgc actacaaacg caaacacgga cacaaaccct tctcgtgtcg 1320  
cctttgcggt aagctttttgg ctgtcaaggg cgattggcga acacatgaga agaattgtgg 1380  
aaaacgttgg gtttgogttt gcggttctga ttttaaacac aaacgttctc ttaaggacca 1440  
tgtaaggcg tttgggtctg gtcattggcc ttatccaact ggtttgtttg aagagcaggc 1500  
15 ttctaattca tctgtctcgg agactttgtt tttttaaatt tgggcatctt tttctttcgc 1560  
ttatgaaata tctattttact ttagaaaaat aataatgtgg tatctaattg ttccaaatta 1620  
ggaacacgaa gtgtaccatt atatttttca tcactacaaa tgttattcag agaaaattat 1680  
cattaattgt ctcgttaaag atagaatagg gtttgaattt atcaaattat aaaaacagat 1740  
caatacaaaa ttgaccatgc atatgcactt gaatattctg atttctttat gatgtaatct 1800  
20 cattcaagaa aagctt 1816

## Patentansprüche

1. Verfahren zur Herstellung einer Pflanze mit veränderter Genexpression, umfassend das  
5 stabile Integrieren einer samenspezifischen regulatorischen Sequenz oder deren Fragment oder  
Derivat und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder  
Derivat funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz in das  
Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen  
Pflanzenzellen oder Pflanzengewebe zu Pflanzen.  
10
2. Verfahren nach Anspruch 1, wobei die Genexpression verstärkt oder verringert ist.
3. Verfahren nach Anspruch 1 oder 2, wobei für die für ein Genprodukt codierende  
15 Nukleinsäuresequenz eine endogene oder exogene Nukleinsäuresequenz verwendet wird.
4. Verfahren nach einem der Ansprüche 1 bis 3, wobei für die für ein Genprodukt  
codierende Nukleinsäuresequenz ausgewählt aus der Gruppe der Gene des  
Phenylpropanoidstoffwechsels, samenspezifischer Gene, samenschalenspezifischer Gene oder  
der Gene des allgemeinen Stoffwechsels verwendet wird.  
20
5. Verfahren nach Anspruch 4, wobei für die Gene des Phenylpropanoidstoffwechsels eine  
Nukleinsäuresequenz verwendet wird ausgewählt aus der Gruppe der Gene für Phenylalanin-  
Ammonium-Lyase, Cinnamat-4-Hydroxylase, 4-Coumarat-CoA-Ligase, Chalcon-Synthase,  
Chalcon-Isomerase, Chalcon-Reduktase, Flavanon 3-Hydroxylase, Flavonoid-3'-Hydroxylase,  
25 Flavonoid-3'-5'-Hydroxylase, Dihydroflavonol-4-Reduktase, Leucoanthocyanidin-Dioxygenase,  
3'-Glucosyltransferase, 5'-Glucosyltransferase und O-Methyl-Transferase.
6. Verfahren nach Anspruch 4, wobei für die samenschalenspezifischen Gene eine  
Nukleinsäuresequenz verwendet wird ausgewählt aus der Gruppe der Gene für  $\alpha$ -Amylase  
30 Inhibitoren, Proteinase-Inhibitoren, Faserproteine und des TT1-Gens gemäß SEQ ID NO:2 und  
SEQ ID NO:4.
7. Verfahren nach Anspruch 4, wobei für die Gene des allgemeinen Stoffwechsels eine  
Nukleinsäuresequenz verwendet wird ausgewählt aus der Gruppe der Gene für ADP-Glucose-  
35 Synthetase, Stärkesynthase, ADP-Glucose-Pyrophosphorylase und Hefe-Invertase.



8. Verfahren nach einem der Ansprüche 1 bis 7, wobei für die samenspezifische regulatorische Sequenz die Nukleinsäuresequenz gemäß SEQ ID NO:1 oder deren Fragment oder Derivat verwendet wird.

5

9. Transformierte Pflanzenzelle oder transformiertes Pflanzengewebe, dadurch gekennzeichnet, daß eine samenspezifische regulatorische Sequenz oder deren Fragment oder Derivat und einer mit der samenspezifischen regulatorischen Sequenz oder deren Fragment oder Derivat funktional verbundenen für ein Genprodukt codierenden Nukleinsäuresequenz stabil in  
10 das Genom der Pflanzenzelle oder des Pflanzengewebes integriert ist.

10. Nukleinsäuresequenz gemäß SEQ ID NO:1 oder deren Fragment oder Derivat oder eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert und für die samenspezifische Expression verantwortlich ist.

15

11. Nukleinsäuresequenz nach Anspruch 10, wobei die hybridisierende Nukleinsäuresequenz unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:1 hybridisiert.

12. Verfahren zur Herstellung von Pflanzen mit verändertem Flavonoidgehalt, umfassend  
20 das stabile Integrieren mindestens der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder einer dazu homologen Nukleinsäuresequenz, oder deren Fragment oder Derivat in das Genom von Pflanzenzellen oder Pflanzengeweben und Regeneration der erhaltenen Pflanzenzellen oder Pflanzengeweben zu Pflanzen.

25

13. Verfahren nach Anspruch 12, wobei die integrierte Nukleinsäuresequenz oder deren Fragment oder Derivat in Sense- oder Antisense-Orientierung zur den Flavonoidgehalt steuernden endogen Nukleinsäuresequenz exprimiert wird.

30

14. Verfahren nach Anspruch 12 oder 13, wobei die Bildung von Flavonoiden durch ein Ribozym, umfassend die integrierte Nukleinsäuresequenz oder deren Fragment oder Derivat unterdrückt wird.

15. Verfahren nach Anspruch 12 oder 13, wobei die Nukleinsäuresequenz oder deren Fragment oder Derivat in den genomischen Bereich des homologen endogenen Gens durch

homologe Rekombination integriert wird.

16. Verfahren nach einem der Ansprüche 12 bis 15, wobei die Nukleinsäuresequenz oder deren Fragment oder Derivat funktional mit einer regulatorischen DNA-Sequenz verbunden ist, die die Expression der integrierter Nukleinsäuresequenz oder deren Fragment oder Derivat steuert.

17. Verfahren nach Anspruch 16, wobei die regulatorische DNA-Sequenz ausgewählt ist aus der Gruppe der Promotoren CaMV 35S-Promotor, PRPI-Promotor, Phaseolin-Promotor, Isoflavon-Reduktase Promotor, ST-LSI Promotor, durch Salizylsäure induzierbarer Promotor, durch Benzenesulfonamid induzierbarer Promotor, durch Tetrazyklin induzierbarer Promotor, durch Abscisinsäure induzierbarer Promotor, durch Ethanol oder Cyclohexanon induzierbarer Promotor, Promotor gemäß SEQ ID NO:1 oder ein samenspezifischer Promotor aus Tabak.

18. Transformierte Pflanzenzelle oder transformiertes Pflanzengewebe, dadurch gekennzeichnet, daß eine den Flavonoidgehalt verändernde Nukleinsäuresequenz oder deren Fragment oder Derivat stabil in das Genom der Pflanzenzelle oder des Pflanzengewebes integriert ist.

19. Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 oder deren Fragment oder Derivat oder eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert und für die Bildung von Flavonoiden verantwortlich ist.

20. Nukleinsäuresequenz nach Anspruch 19, wobei die hybridisierende Nukleinsäuresequenz unter stringenten Bedingungen mit der Nukleinsäuresequenz gemäß SEQ ID NO:2 oder 4 hybridisiert.

21. Aminosäuresequenz wie in SEQ ID NO:3 aufgeführt.

22. Pflanzenzelle oder Pflanzengewebe nach Anspruch 9 oder 18, regenerierbar zu einer samenproduzierenden Pflanze.

23. Pflanze, erhältlich nach einem der Ansprüche 1 bis 8 oder 12 bis 17.



24. Saatgut, erhalten von Pflanzen nach Anspruch 23.
25. Vektor, umfassend eine Nukleinsäuresequenz nach einem der Ansprüche 10 oder 11 oder 19 oder 20.

FIG. 1 A



1 actagttgaccacatgaactaaacttcttggacaatcatcaatggacaca  
 51 tgttagctttgatttgctgtgaatttggtttatctctcagtataattatc  
 101 actttcttggtttatgcttacaatatattttatgggttagagttttgttt  
 151 acgattttggatttaattggataaagattagggattgagggtttgagttta  
 201 gggtaaggaaatttaggccttagtgtagagtcctcaagggtttaagggttac  
 251 acaccacaaaccatttgcttgtgtcaacaacattgtatcatattttcaaa  
 301 aaaattttgttgaaggaccttgattgatataataaaagcgaactgtttg  
 351 gataagtttatgtggacaatatattggatacataattagaaacatagt  
 401 ttaatatctgatatttggtgggaatatataatactacttaggtttaaata  
 451 tatagtatttcatatgatgcgaactgtttggataagtttacgtggacaat  
 501 atatatattgatacataattaggaacatagtttaatatattgatatttggtg  
 551 ggaatatataattctacttacgcttaaatatttttatttgaattaaagca  
 601 tttcatataatgtgaactgtttgaatatgtttacatggacaatatatatt  
 651 ggatacataattaggaacatagtttaatatctgatatttggtggaaatat  
 701 ataataattagtttaagcttaaatatttttatttgatataatatttgactta  
 751 aacatttttatttgattaaactaaattttaacagatcttaccattaattt  
 801 ttaacttggtatctctatctaattgtcacgtatattgttttttagtaattg  
 851 gcaacaaaattaatttatctctctgttttttttcttctcacctttataag  
 901 ggtaaaatggtcataaaatcagtaaaaaagggtggaagggtgccactccc  
 951 tcaaaagtgtcataaacgtccaaactttctccataaatgccttatttttg  
 1001 aacattccatataagattataaacttattataggttataacttattatagtt  
 1051 acgttaattatatgaatttctatttagttatcacacaatcaaaTATTTTAA  
 1101 TCACAAaaatttattaaacattttatatgtggtagtataatgcaataaca  
 1151 tattatatgtggtggcataatgcaacaacatattatttgcctacgaatct  
 1201 cctttatttttctgtttatgtaacaacagtaaaacggattgtttagcttga  
 1251 tattctataattataaataatctaaagttatttttgtaaattatttttttct  
 1301 caaattggataaccaatcatagagttggtatttttttTTTGTGTTAAAT  
 1351 Atatatatactgaatatcgagttattgtgcatgacaatttatatggcgga  
 1401 cgagtttaaactcgacattaataacaattaaaatattattaatctaatact  
 1451 taaatactggttaaatacacaatttattattcttaataccacataattaaa  
 1501 catatctaacttttactgattcaataaagattgtgtgaaacaaaagttgt  
 1551 cttgcaaagaattaatattgtacatagattttgttctggttagctagtact  
 1601 aaaatccattaataaaaactaatacgggtatctttattgatcatgtaacatg  
 1651 aattattcatgtatatacaattgaccctattaattttgcataaaattcaa  
 1701 cttggcaaattcattgattttgtaaaccgttaattctgctaatttcacaa  
 1751 ttctctgtacgctaaaaatttatgcgtattatcgtattgatatgcaaatt  
 1801 atcgaagaatttatagttttatatagtagaaatgaaggattttgcaaaac  
 1851 gagttctaactgaaataacactaattaattaattagagtttgaacctac  
 1901 agagattcgacttgatccacttgaaaaattcatttactctactaatttg  
 1951 ttactccatggaccatgattatgctattctgtaggactctaacaactgac  
 2001 ttgacacaatctctttcgtgaacaataatgggttatattttttgttttg  
 2051 ttttttcggacaaattagccacgttgcttttagaccattttgtagttctta

FIG. 1 B



2101 tcttgaatcaaagtctcagctaaaaaaaaAAAAACGCTTAAatccac  
2151 tagctagactacgactacgttggttaaagtgttttTTTTTAAATACAAtac  
2201 aTTGAAGTTAAATATttgaataaagaaaatctaatacagcatgtatacagt  
2251 atattagaagtaataacttgatcagaaaaataacatacaataataaaataa  
2301 taaaaaaattatgttagtttttgggaatattataatttctactttcaatca  
2351 aaataactaaaagaaataaaatcttcacacatagtGGTAATAATTGGCTa  
2401 gtatgaatattgaattgtggagacccggcataatatttgactaggcagaa  
2451 attattgatatgtactaagttaataaccttgcaaagaaattcttttagtg  
2501 aaacgtgtacatttgtaaaaacagatttaacactaaatcttgacttgat  
2551 atactattaattatttcttttctcttatttggtatgtcaaatctagtgttt  
2601 acaaaaccagaggtgttgaccgttagagagagaattaaacaacttacata  
2651 catacaaaacataacccaaaaaataataaatgcatcttccataataa  
2701 taataatatgaattcaacattagcattcatttcattacccaaatccgaaa  
2751 tttcattgattaaaaattaatacaattgtattgtagaaaagctaaaagctt  
2801 acgtaaatgccaaagatagtcaaaaccctgcaatgacaaagttgccaaaa  
2851 tcttgaagagtttggtccacaaaatttaagggtcttggttttccactcta  
2901 tttataggcaaagagatgagacagagaagattaaattacttcttaacaaa  
2951 ggttggttttcaactcaaccacatgcattctcaagtgtctgctcctcacatt  
3001 cccaagattcccatttactcacttctctatttggtacgtaagtccacaca  
3051 atatgattctaaattattttacacattattcgttttgttcacacttgctt  
3101 tcgactttcgtaaacctatatagttcatccaatattattcggtaaattcg  
3151 atatttatcaatctttattctcgtaggttaaaggagacgattgatacgtg  
3201 ggatctacttacgtatctgcatgattattagttataaaagttattgcaaa  
3251 cattaaattactttcatagagagcaatcattatattaaggtaatttaatt  
3301 ttattatatatagtcgaagatttaaggaataaagaaaagattctcaaaac  
3351 atttcatctctctccaacaactattcaccacattcaATG



3387 ATGGAGTCACCACC  
3401 ACTATACGAGATATCCTCAAGCTCTTCTCTGAAAAACCTAGACACCATT  
3451 TCCAATCCCTTGATCTCTTCCCTAACCTCAACCAAACTCTTGTATCAAC  
3501 AATACCCTAATTGAGCCTTTACCGCTTATTGATCGCATAACTTGAAGTC  
3551 AAACCTAGACCTAAACCCTAATCCCTTGATGCGGAAGAAGGAGAGCAAG  
3601 AGGAGGAAGAAGAAGAAGAAGACCGTGAAGTGGACGTGGACTTACAC  
3651 ATCGGCCTTCCTGGTTTTGGTAAACCAAGCAATGATGCTAAACAGCTGAA  
3701 GAAGAGAAATGGGAAGGAGATCGCCACATATGACGCCGAAAAGGCATCG  
3751 AGAATGAACTTTCCGAAAGGCATACTGGATCCCGGCGCCGGAGCAAATT  
3801 CTCATAGGGTTCACTCATTTTTCTTGCCATGTATGCTTCAAGACATTCAA  
3851 TCGCTACAACAATCTTCAGgtacgagtcaatatatctcatgcgcatgtgt  
3901 tttccatgcacaaacatatataataaattcatcttatagagttatatctc  
3951 cggatctaattgttatgagttttattcatatctatatatacatatatata  
4001 tatatatatatatatatatataaattctgaatttttttgataaaagc  
4051 taaacaaccaggattttaatagatgatttacctttggatcttattatacaa  
4101 tttacaaatttaatacaagtcaactaatcgtgatttaattacttttttttg  
4151 taagaagagttggtaatatatatatttttatggtaattgttttcatgaaaata  
4201 attcatcacaaactctttacattttatttaattgccttaactaaagctgaatt  
4251 cgaaaaagttgaaataaattatctactaagatttgattgactatagtttt  
4301 taatagtttttcttttctcatatatatatattatcatagtagtcaaaacattt  
4351 gattcaaacttaaaatacacagatttcttgatgaaacattactatgctcg  
4401 gtcaataatatgatttttaaggaaccatgttatttcatttttattacttaag  
4451 gaaacctttttgttttttgttgactctaaatattatgaatatagATGCAC  
4501 ATGTGGGGACATGGTTCACAATACAGGAAAGGACCGGAGTCACTGAAAGG  
4551 CACACAGCCACGAGCCATGTTAGGGATCCCTTGTTACTGCTGCGTTGAAG  
4601 GGTGCAGGAACCACATTGACCATCCTCGTTCCAAGCCACTGAAAGACTTT  
4651 AGGACGCTCCAAACGCACTACAAACGCAAACACGGACACAAACCCTTCTC  
4701 GTGTCGCCTTTGCGGTAAGCTTTTGGCTGTCAAGGGCGATTGGCGAACAC  
4751 ATGAGAAGAATTGTGGAAAACGTTGGGTTTGCCTTTGCGGTTCTGATTTT  
4801 AAACACAAACGTTCTCTTAAGGACCATGTTAAGGCGTTTGGGTCTGGTCA  
4851 TGGGCCTTATCCAACCTGGTTTGTGTTGAAGAGCAGGCTTCTAATTCATCTG  
4901 TCTCCGAGACTTTGTTTTTTTAAatttgggcatctttttctttcgcttat  
4951 gaaatatctattttacttttagaaaaataaataatgtggtatctaattgttcc  
5001 aaattaggaacacgaagtgtaccattatatttttcatcactacaaatggt  
5051 attcagagaaaaattatcattaattgtctcgttaaagatagaatagggttt  
5101 gaatttatcaaataattaaaaacagatcaatacaaaattgaccatgcatat  
5151 gcacttgaatattctgatttctttatgatgtaattctcattcaagaaaagc  
5201 tt

FIG: 3

ATGGAGTCACCACCACTATACGAGATATCCTCAAGCTCTTCTTCTGAAAAACCTAGACAC  
 1 -----+-----+-----+-----+-----+ 60  
 M E S P P L Y E I S S S S S S E K P R H -  
 CATTTCGAATCCCTTGATCTCTTCCCTAACCTCAACCAAACTCTTGTATCAACAATACC  
 61 -----+-----+-----+-----+-----+ 120  
 H F Q S L D L F P N L N Q N S C I N N T -  
 CTAATTGAGCCTTTACCGCTTATTGATCGCATAACTTGAACCTCAAACCTAGACCTAAAC  
 121 -----+-----+-----+-----+-----+ 180  
 L I E P L P L I D R I N L N S N L D L N -  
 CCTAATCCCTTGATGCGGAAGAAGGAGAGCAAGAGGAGGAAGAAGAAGAAGAAGAC  
 181 -----+-----+-----+-----+-----+ 240  
 P N P L Y A E E G E Q E E E E E E E E D -  
 CGTGAAGTGGACGTGGACTTACACATCGGCCTTCTGGTTTTGGTAAACCAAGCAATGAT  
 241 -----+-----+-----+-----+-----+ 300  
 R E V D V D L H I G L P G F G K P S N D -  
 GCTAAACAGCTGAAGAAGAGAAATGGGAAGGAGATCGCCACATATGACGCCGGAAGGC  
 301 -----+-----+-----+-----+-----+ 360  
 A K Q L K K R N G K E I A T Y D A G K G -  
 ATCGAGAATGAACCTTCCGGAAGGCATACTGGATCCCGGCGCCGGAGCAAATTCTCATA  
 361 -----+-----+-----+-----+-----+ 420  
 I E N E L S G K A Y W I P A P E Q I L I -  
 GGGTTCACTCATTTTTCTTGCCATGTATGCTTCAAGACATTCAATCGCTACAACAATCTT  
 421 -----+-----+-----+-----+-----+ 480  
 G F T H F S C H V C F K T F N R Y N N L -  
 CAGATGCACATGTGGGGACATGGTTACAATACAGGAAAGGACCGGAGTCACTGAAAGGC  
 481 -----+-----+-----+-----+-----+ 540  
 Q M H M W G H G S Q Y R K G P E S L K G -  
 ACACAGCCACGAGCCATGTTAGGGATCCCTTGTACTGCTGCGTTGAAGGGTGCAGGAAC  
 541 -----+-----+-----+-----+-----+ 600  
 T Q P R A M L G I P C Y C C V E G C R N -  
 CACATTGACCATCCTCGTTCCAAGCCACTGAAAGACTTTAGGACGCTCCAAACGCACTAC  
 601 -----+-----+-----+-----+-----+ 660  
 H I D H P R S K P L K D F R T L Q T H Y -  
 AAACGCAAACACGGACACAAACCCTTCTCGTGTGCGCTTTGCGGTAAGCTTTTGGCTGTC  
 661 -----+-----+-----+-----+-----+ 720  
 K R K H G H K P F S C R L C G K L L A V -  
 AAGGGCGATTGGCGAACACATGAGAAGAATTGTGGAAAACGTTGGGTTTGGCTTTGCGGT  
 721 -----+-----+-----+-----+-----+ 780  
 K G D W R T H E K N C G K R W V C V C G -  
 TCTGATTTTAAACACAAACGTTCTCTTAAGGACCATGTTAAGGCGTTTGGGTCTGGTCAT  
 781 -----+-----+-----+-----+-----+ 840  
 S D F K H K R S L K D H V K A F G S G H -  
 GGGCCTTATCCAACCTGGTTTGTGTTGAAGAGCAGGCTTCTAATTCATCTGTCTCCGAGACT  
 841 -----+-----+-----+-----+-----+ 900  
 G P Y P T G L F E E Q A S N S S V S E T -  
 TTGTTTTTTTAA  
 901 -----+----- 912  
 L F F \* -

FIG. 4

	1				50
AtTT1	.....	.....	.....	.....	.....
AtAL049660	.....	.....	.....	.....	.....
AtAB025629	MLFSTVLSHR	TLYLTCNPNT	LIHSYTHPHI	HAYLAFTGFL	TQLHHLEISC
AtAC006085.9	.....	.....	.....	.....	.....
Hv234704	.....	.....	.....	.....	.....
Consensus	.....	.....	.....	.....	.....
	51				100
AtTT1	.....	.....	.....	MESPP	LYEISSSSSS
AtAL049660	.....MTPD	YSNFFTDWFK	SNPFHH..YP	NSTNPSPPH	LPPVTPPSSS
AtAB025629	LLLLFFSLSS	LLKLMADPDC	IFRNGYVDYY	NYSFNYATSL	SRIYNSHDSF
AtAC006085.9	.....	MSNPAC	SNLFNNGCDH	N.SFNYSTSL	SYIYNSHGSY
Hv234704	.....	.....	.....	.....	.....
Consensus	.....	.....	.....	.....	.....S.
	101				150
AtTT1	EKPRHHFQSL	DLFPNLNQNS	CINNTLIEPL	PLIDRINLNS	NLDLNPNP..
AtAL049660	FFFFQSGD..	LRRPPPPPTP	PPSPPLREAL	PLLSLSPANK	QQDHHNH.D
AtAB025629	YYPHQTTNPN	INE.NPNLTS	PDSPPLREAL	PLLSLSPIHK	HQEPTANHHE
AtAC006085.9	YYSNTTNPNY	INHHTTTSTS	PNSPPLREAL	PLLSLSPI.R	HQEQQDQH..
Hv234704	.....	.....	.....	.....	.....
Consensus	.....	.....	.....l.e.l	pl.....	.....
	151				200
AtTT1	.LYAEEGEQE	EEEEEEEDRE	.....	VDVDLHIG	LPGFG.....
AtAL049660	HLIQEPPSTS	MDVDYDHHHQ	DDHNLDDDD	HDVTVALHIG	LPSPSAQEMA
AtAB025629	YYFMETTETS	SNSNFLDQCQ	DSYR.....	HDVTVDLHLG	LPNLGDGG..
AtAC006085.9	.YFMDTHQIS	S.SNFLDDPL	.....	VTVDLHLG	LPNYGVGE..
Hv234704	.....	.....	.....	.....	.....
Consensus	.....	.....	.....	v.v.lh.g	lp.....
	201				250
AtTT1	.....KPSND	AKQLKCRNGK	EIATYDAGKG	IENELSGKA.	.....
AtAL049660	SLLMSSSSSS	SSRTTHHHED	MNHKKDLDE	YSHGAVGGGE	DDDEDSVGGD
AtAB025629	.....SSSD	VULDSTDHQE	GHHDDHQQDQ	LEVMTAS...	DHDDEHGGGLQ
AtAC006085.9	.....SIRSN	IAPDATTDEQ	...DQDHDRG	VEVTVESHL	DDDDHHGDLH
Hv234704	.....	.....	.....	.....	.....
Consensus	.....	.....	.....	.....	.....
	251				300
AtTT1	.....	YWIPAPEQ	ILIGFTHFSC	HVCFKTFNRY	NNLQMHMWGH
AtAL049660	GGCRISRLNK	GQYWIPTPSQ	ILIGPTQFSC	PVCFKTFNRY	NNMQMHMWGH
AtAB025629	RGNHLHH...	FWIPTPSQ	ILMGPTQFSC	PLCFKTFNRY	NNMQMHMWGH
AtAC006085.9	RG...HH...	YWIPTPSQ	ILIGPTQFTC	PLCFKTFNRY	NNMQMHMWGH
Hv234704	.....	.....	.....	.....	...QMHWGH
Consensus	.....	wip.p.q	il.g.t.f.c	..cfktnry	nn.QMHWGH
	301				350
AtTT1	GSQYRKGPES	LKGTQPRAML	GIPCYCCVEG	CRNHIDHPRS	KPLKDFRTLQ
AtAL049660	GSQYRKGPES	LRGTQPTGML	RLPCYCCAPG	CRNNIDHPRA	KPLKDFRTLQ
AtAB025629	GSQYRKGPES	LRGTQPTAML	KLPCYCCAPG	CKNNIDHPRA	RPLKDFRTLQ
AtAC006085.9	GSQYRKGPES	LRGTQPTGML	RLPCFCCAPG	CKNNIDHPRA	KPLKDFRTLQ
Hv234704	GREYRKGPES	LKGTQTVAL	KVPCYCAA.G	CRNSVSHPR	RPLKDFRT..
Consensus	Gs#YRKGPES	LkGTQp.a\$L	k.PCYCca.G	CrN.!dHPRA	rPLKDFRTLq
	351				400
AtTT1	THYKRKHGK	PFSCRLCGKL	LAVKGDWRTH	EKNCGKRWVC	VCGSDFKHKR
AtAL049660	THYKRKHGK	PFMCRCGKA	FAVRGDWRTH	EKNCGKLWYC	ICGSDFKHKR
AtAB025629	THYKRKHGVR	PFACRRCGKA	FAVKGDWRTH	EKNCGKLWYC	SCGSDFKHKR
AtAC006085.9	THYKRKHGSK	PFACRMCGKA	FAVKGDWRTH	EKNCGKLWYC	SCGSDFKHKR
Hv234704	.....	.....	.....	.....	.....
Consensus	thykrkhg..	pf.cr.cgk.	av.gdwrth	ekncgk.w.c	.cg sdfkhkr
	401				448
AtTT1	SLKDHVKAFG	SGHGPYPTG.	.LFEEQASNS	SVSETLFF..	.....
AtAL049660	SLKDHVKAFG	NGHGAYGID.	.GFDEED..E	PASEVEQLDN	DHESMQSK
AtAB025629	SLKDHVKAFG	NGHVPC....	CGIDHEEE.E	AASDVEQQE.	.....
AtAC006085.9	SLKDHVKAFG	NGHVPCGIDS	FGGDHEDYYD	AASDIEQ...	.....
Hv234704	.....	.....	.....	.....	.....
Consensus	slkdh.kafg	.gh.....	.....	..s.....	.....



FIG. 5

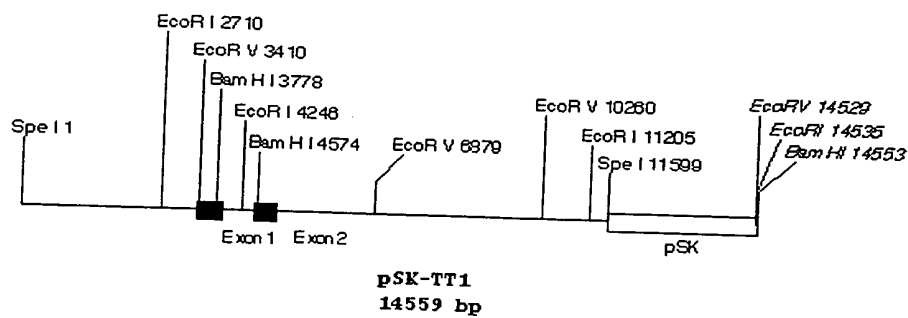


FIG. 6

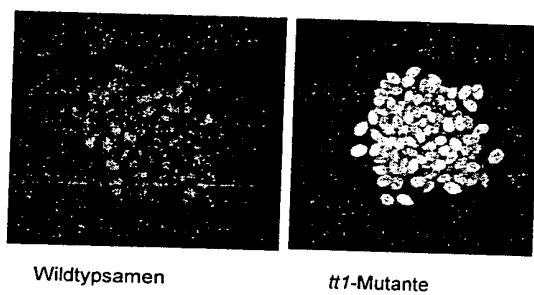
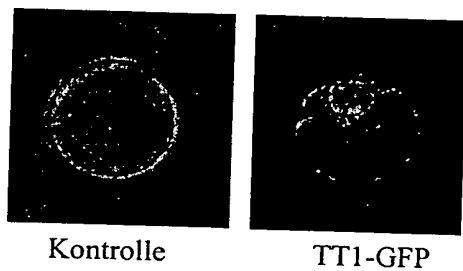


FIG. 7





---